# Garbage Detection using YOLO Algorithm for Urban Management in Bangkok

MATHUROS PANMUANG[1], CHONNIKARN RODMORN[2*]
[1]Department of Educational Technology and Communications, Faculty of Technical Education,
Rajamangala University of Technology Thanyaburi,
Pathum Thani 12110,
THAILAND

[2]Department of Applied Statistics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok,
Bangkok 10800,
THAILAND

*Corresponding Author*

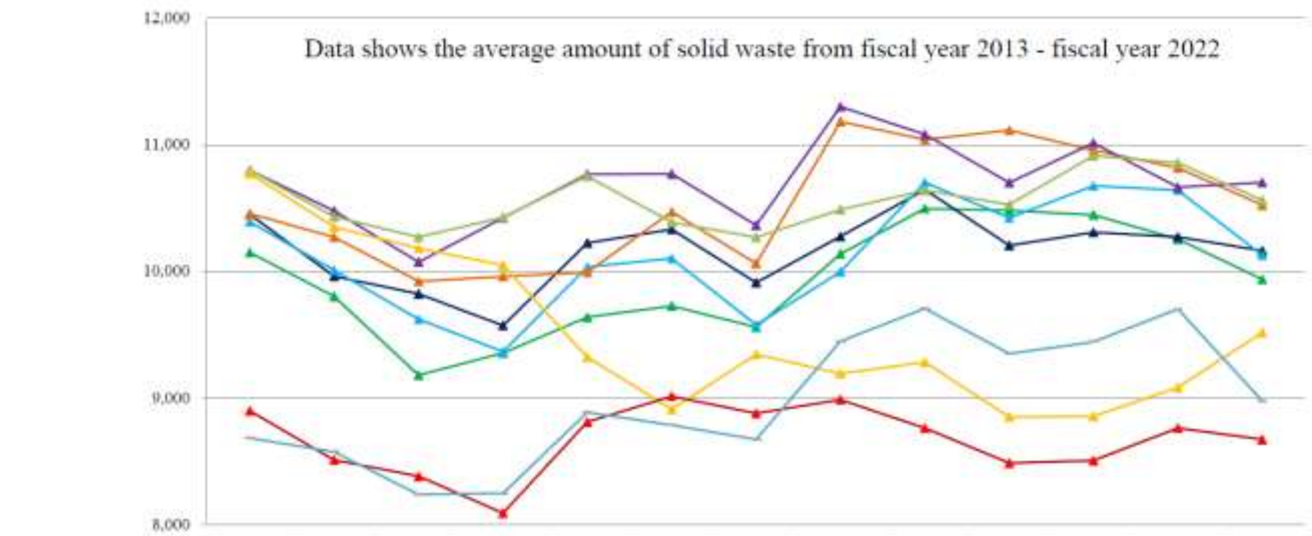
*Abstract:* - Garbage problems in urban areas are becoming more serious as the population increases, resulting in community garbage, including Bangkok, the capital of Thailand, being affected by pollution from rotten waste. Therefore, this research aims to apply deep learning technology to detect images from CCTV cameras in urban areas of Bangkok by using YOLO to detect images from CCTV cameras in urban areas of Bangkok, using YOLO to detect 1,383 images of overflowing garbage bins, classified into 2 classes: garbage class and bin class. YOLO in each version was compared, consisting of YOLOv5n, YOLOv6n, YOLOv7, and YOLOv8n. The comparison results showed that YOLOv5n was able to classify classes with an accuracy of 94.50%, followed by YOLOv8n at 93.80%, YOLOv6n at 71.60%, and YOLOv7 at 24.60%, respectively. The results from this research can be applied to develop a mobile or web application to notify of overflowing garbage bins by integrating with CCTV cameras installed in communities to monitor garbage that is overflowing or outside the bin and notify relevant agencies or the locals. This will allow for faster and more efficient waste management.

*Key-Words:* - Garbage detection, Overflowing garbage bins, YOLO, Deep Learning, Machine Learning, Image Processing.

Received: March 19, 2023. Revised: October 26, 2023. Accepted: December 27, 2023. Published: February 27, 2024.

## 1 Introduction

The world's population is growing and is likely to reach 10 billion in 2050 from the current level of 7.6 billion, causing urban expansion, increasing rapidly as the human population increases. As a result of urban expansion, environmental problems follow, [1]. Environmental problems have become a problem in the world and Thailand today as people neglect to take care of proper management of natural resources and the environment, causing impacts on the ecosystem and human life, [2]. Key environmental problems are garbage or solid waste. Currently, the amount of solid waste or waste is increasing every year which is caused of an increase in population, and economic and industrial expansion, which is considered an important problem for the community that must be managed and solved in terms of solid waste, sewage, and toxins that contaminate water sources, soil, and air, [3]. Foreign studies have found that developing countries are unable to manage solid waste sustainably, resulting in severe environmental degradation, [4]. At present, it is found that every urban community in Thailand faces problems with the density of houses and the large number of businesses and shops, but with limited space and natural resources remaining the same as well as an increasing population, which results in the problem of community waste, [5]. Bangkok, the capital of Thailand, is the center of prosperity in every aspect, both economically and socially. Bangkok is therefore a city with a problem with the amount of solid waste. The Office of the Environment's annual work plan report for 2016 shows that each year Bangkok has a continuous increase in the amount of solid waste and found that in 2015, Bangkok had a high amount of solid waste of 10,167 tons/day and, [6].




| Fiscal year | OCT | NOV | DEC | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2557 | 10,151.81 | 9,805.02 | 9,182.45 | 9,358.70 | 9,640.07 | 9,729.90 | 9,561.98 | 10,141.33 | 10,499.62 | 10,483.19 | 10,450.34 | 10,255.97 | 9,939.80 |
| 2558 | 10,453.62 | 9,963.28 | 9,824.47 | 9,574.51 | 10,226.50 | 10,334.24 | 9,914.28 | 10,277.78 | 10,648.68 | 10,208.04 | 10,311.58 | 10,275.22 | 10,166.84 |
| 2559 | 10,392.96 | 10,010.45 | 9,625.62 | 9,367.80 | 10,037.67 | 10,104.55 | 9,578.92 | 9,998.77 | 10,705.93 | 10,424.22 | 10,679.16 | 10,644.10 | 10,130.22 |
| 2560 | 10,458.37 | 10,273.48 | 9,923.22 | 9,960.95 | 9,996.12 | 10,472.60 | 10,065.41 | 11,185.71 | 11,042.45 | 11,117.11 | 10,958.91 | 10,820.40 | 10,526.92 |
| 2561 | 10,800.73 | 10,481.41 | 10,076.60 | 10,424.33 | 10,769.29 | 10,772.94 | 10,367.71 | 11,304.00 | 11,084.05 | 10,703.51 | 11,016.64 | 10,667.66 | 10,705.82 |
| 2562 | 10,800.01 | 10,433.93 | 10,273.97 | 10,429.21 | 10,753.95 | 10,389.77 | 10,271.98 | 10,491.25 | 10,642.54 | 10,528.30 | 10,916.43 | 10,859.10 | 10,564.48 |
| 2563 | 10,775.38 | 10,351.98 | 10,185.50 | 10,050.81 | 9,326.06 | 8,912.45 | 9,347.35 | 9,194.71 | 9,286.21 | 8,852.65 | 8,856.53 | 9,085.31 | 9,519.81 |
| 2564 | 8,899.92 | 8,513.25 | 8,384.28 | 8,093.35 | 8,812.41 | 9,016.26 | 8,881.98 | 8,989.12 | 8,765.11 | 8,488.35 | 8,508.15 | 8,765.21 | 8,674.73 |
| 2565 | 8,686.26 | 8,576.42 | 8,240.34 | 8,250.67 | 8,887.89 | 8,788.64 | 8,675.24 | 9,450.55 | 9,709.05 | 9,352.14 | 9,445.40 | 9,704.97 | 8,979.37 |

Fig. 1: Shows the average amount of solid waste from fiscal year 2013 - fiscal year 2022, [7]

The amount of waste found from a survey by the Environment Agency from fiscal year 2013 - fiscal year 2022 found that the amount of waste each year is similar or may decrease in certain years, but still considered to be a large amount (Figure 1).

At present, the Environment Agency has given serious importance to waste management and campaigns to make people aware of their responsibilities in sorting and disposing of waste at designated points. However, it still faces complaints from citizens about the amount of residual waste that gives off a bad odor and creates an unattractive environment and the waste disposal that does not comply with regulations, such as failure to dispose of waste timely or in the right place. Currently, Bangkok has installed more than 60,000 CCTV cameras in the area, divided into CCTV cameras to monitor traffic conditions and CCTV cameras for safety. Therefore, to integrate, utilizing existing CCTV cameras for waste management will make waste management more efficient and create greater satisfaction for citizens.

Therefore, in this research, image processing techniques were applied to help manage waste by analyzing images from CCTV cameras and using the YOLO model for processing from deep learning to classify the images into 2 classes: overflowing garbage bin and bin class. The results obtained from this research will allow Bangkok to develop further into a mobile application to notify relevant agencies or people in the area to know whether the waste is overflowing. It is also a process that allows the public, government, and private sectors to participate in sustainable waste management.

## 2 Related Work

The 2016 algorithm known as YOLO (You Only Look Once) offers quick speed and excellent accuracy. The input image is divided into a grid, and the likelihood of the bounding box and related objects appearing in each grid cell is predicted. It performs bounding box estimation and object categorization using these predictions. Because of its excellent accuracy and quick speed, it has been used extensively in the field of table recognition recently, [8]. The YOLO architecture is a single-stage object detector with three main parts: the first part is the backbone, which is responsible for extracting image features; the second part is the model neck, a layer between the backbone and the head that is used to extract different feature maps of different states within the backbone, which helps in dealing with images with different image scales such as the feature pyramid network; and the last part, the model head, is responsible for detecting objects and is used to predict images in the bounding box and calculate confidence scores. Currently, YOLO is so popular that it has been developed in several versions; until now in 2023, it has developed into YOLOv8.

There are differences in the bounding boxes between the two dataset formats. The network of YOLO is made up of two fully linked layers after twenty-four convolutional layers. The following is how the YOLO framework works. There are n × n grids inside the image. Next, each grid is subjected to image localization and classification. The bounding boxes and the associated class probabilities are predicted by YOLO. Labeled data is needed to train the model. A one-n-dimensional vector called $y$ is present in each grid cell. The $y$ value is an 8-dimensional vector made up of the following if the image is divided into a 3 × 3 grid and three classes need to be classified, [9].

$$y = [p_c, b_x, b_y, b_h, b_w, c_1, c_2, c_3] \tag{1}$$

Where $p_c$ is a number that indicates if an object is present in that grid cell and might be either 0 or 1. The bounding box's coordinates are described by $b(x,\ y,\ h,$ and $w)$, and the existence of a certain class in a given grid cell is indicated by the binary values $c_1$, $c_2$, and $c_3$.

For research that applies YOLO to object detection, such as employing YOLOv4 for real-time vehicle detection with improvement for increased efficiency, refer to, [10]. The YOLOv4-tiny model combined with Hard Negative Sample Mining (HNEM) uses the CSPBlock module by using the CSPDarknet-53 as the backbone to improve the detection rate of occluded vehicles. The experiment results found that improving the efficiency of real-time vehicle detection becomes more accurate compared to the research, [11], that used YOLOv5 to categorize the parts according to their lateral and head shapes. The creation of an image acquisition platform with two mounted cameras and an appropriate lighting system to produce high-quality photos is another element of this work. The suggested deep learning (YOLOv5) method has demonstrated encouraging results with a mAP@0.5 of 0.996 for component classification, despite the difficulties connected with such systems. In contrast, the suggested image processing method produced a maximum inaccuracy of 0.05 mm for pitch computation and 100% accuracy for standard-size assignments. In addition, there is an introduction to the utilization of YOLOv6 for transferring learning to a real-time object detection model, [12]. Another crucial component of this work is the suggested model's ability to recognize every object in a scene—indoor and outdoor—and to alert the user to close and distant things via voice output. The Google Text-to-speech (gTTs) library is used to get the audio response. Following 30%, 40%, and 50% pruning of the YOLOv6 baseline model, the optimized YOLOv6 framework achieves 37.8% greater average accuracy (AP) at 1235 frames per second (FPS).

Meanwhile, YOLOv7 is employed to introduce a novel technique for detecting dragon fruit. This technique goes beyond merely locating the fruit; it also identifies the endpoints at its head and root, [13]. YOLOv7 outperforms specific previous models in this regard. Furthermore, the study applied YOLOv8 to detect waste visually, leveraging the efficiency of YOLOv8, the latest object detection model in the YOLO series, specifically for automated waste sorting, [14]. YOLOv8 is used for automated waste sorting to improve safety and efficiency in waste treatment procedures. The outcomes show that YOLOv8 is an effective tool for enhancing waste management procedures since it outperforms state-of-the-art waste detection and classification algorithms.

However, research on the use of YOLO for identifying different things was also conducted, and it was discovered that deep learning techniques are used for garbage detection. The study also creates a garbage image categorization system based on deep learning, [15]. Its primary goals are to compare deep learning neural network models, identify the best classifier, create online apps, and implement neural networks. According to the results, inceptionResnetv2 has an 89% detection accuracy and a 0.8 loss value. It performs better in terms of detection when compared to the two migration models mentioned above.

Based on a review of related research, the researcher applies YOLO for real-time CCTV image detection to sort out overflowing garbage bins in Bangkok by comparing the YOLO architecture in each version including YOLOv5n, YOLOv6n, YOLOv7, and YOLOv8n to get the model with the best accuracy.

# 3 Methodology

## 3.1 Data Collection and Preparation of Datasets

The image datasets from the internet and video images of Bangkok's public garbage dumps were used for testing in this study. Then, the images were enhanced using three image augmentation techniques: horizontal image inversion, brightness adjustment, and 90-degree rotation. This produced 1,383 images, which were then labeled into two classes using the Roboflow tool: bin images and garbage images. Later, the data set was divided into 3 parts: training with 1,105 images, validation with 139 images, and testing with 139 images, or equivalent to 80:10:10. An example of the data set is shown in Figure 2.

Fig. 2: Example of data set

## 3.2 Model Training

The images were resized to the same size, 640×640 pixels, in 3 channels. Then, the images were trained with the validation dataset to measure the performance of the model and set the threshold equal to 0.25, the IoU (Intersection over union) value equal to 0.50, trained for 100 epochs. In each model, the researcher selected the model with the smallest size to reduce the training time as when testing the model, it was found that the smallest size is sufficient for classification. Therefore, the models used include YOLOv5n, YOLOv6n, YOLOv7, and YOLOv8n.

YOLOv5n: The most recent network model in the YOLOv5 series is the YOLOv5n. On the one hand, the YOLOv5n network model features fast reasoning speed and good detection accuracy. In contrast, the YOLOv5n network model's weight file is very small—roughly 75% less than YOLOv5s—meaning that YOLOv5n is well suited for deployment to embedded devices for real-time detection because automatic driving lane curvature target identification relies heavily on the precision, real-timeliness, and lightweight nature of the model, [16].

YOLOv6n: YOLOv6 through the development of a parameterizable network structure based on RepVGG-EfficientRep and Rep-PAN, as well as improvements to the backbone network, neck, detection head, and training method. Therefore, YOLOv6n reduces more parameters and processing, but the average recognition accuracy is significantly lower as a result. Regarding models operating on devices with high-performance computing, [17].

YOLOv7: Additionally, YOLOv7 makes use of the idea of deep supervision. To guide the weight of the external network, it provides an additional auxiliary head structure in the middle network layer. These mono-modality object recognition approaches achieve performance as well as real-time inference. However, these object detection models only use one stream. Consequently, the advantages of each stream—such as precise edges and suitable illumination in infrared images and the object's color and detail information in red-and-green images—cannot be utilized by these models. Enhancing object detection performance across all streams requires appropriate feature exploitation, [18].

YOLOv8n: The four components of YOLOv8n are the input, backbone, neck, and head. The C2f and CBS modules, which extracted features from the input image, made up the majority of the Backbone. The C2f module allowed for the gathering of rich gradient information while being lightweight. The neck implemented feature pyramid networks and path aggregation network topologies, which improved the network's capacity for feature fusion by merging features of various scales. Three distinct scale detection branches were present in the head, and non-maximum suppression (NMS) was used to determine the ideal detection box, [19].

### 3.3 Model Performance Evaluation

The performance of the model is measured by determining the precision, recall, and mean average precision (mAP), which are values obtained from the testing method to determine the predictive value of the accuracy of the data, [20], as detailed below:

1) Precision is a measure of the accuracy of the prediction. If the value is large, it means that the evaluation method has high accuracy as shown in Equation 2.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

2) Recall is a measure of the accuracy of the model by considering each class separately as shown in Equation 3.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

3) Mean Average Precision (mAP) is the average of the Precision and Recall of objects in the image as shown in Equation 4.

$$mAP = \frac{1}{N}\sum_{N}^{i=1} AP_i \tag{4}$$

TP (True positive) is the positive correct prediction value and TN (True negative) is the negative correct prediction value, while FP (False positive) is the positive false prediction value and FN (False negative) is the negative false prediction value.

## 4 Result and Discussion

To compare each version of YOLO, the resources of the computer are determined on the same basis:

- CPU: Intel(R) Core(TM) i7-9700 CPU@3.00GHz
- GPU: NVIDIA® GeForce® GTX 1660 Ti
- RAM: 16 GB
- Storage: 220 GB SSD.
- Operating System: Windows 11 Home Single Language

We also used Google Colab with Tesla T4 GPU to train the neural networks, and 100 epochs were trained for every model, and hyperparameters were configured to default. However, for YOLOv7, after training with 100 epochs, it was found that the performance value was quite low (unable to recognize the class of the test data set). Therefore, we increased the number of epochs from 100 to 200, and the results of increasing the number of trains showed that the efficiency increased. The comparison results are shown as follows:

Table 1. Model performance comparison results

| Model | Precision | Recall | mAP:0.5 | mAP:0.95 |
|---|---|---|---|---|
| YOLOv5n | 0.938 | 0.890 | 0.945 | 0.633 |
| YOLOv6n | 0.717 | 0.617 | 0.716 | 0.370 |
| YOLOv7 | 0.340 | 0.319 | 0.246 | 0.082 |
| YOLOv7 (200 epochs) | 0.893 | 0.758 | 0.868 | 0.466 |
| YOLOv8n | 0.941 | 0.876 | 0.938 | 0.707 |

From Table 1, the comparison of model performance results showed that YOLOv5n had the highest mAP:0.5 value; that is, the model was able to classify classes with an accuracy of 94.50%, which when compared to YOLOv8n, the value was not much different. If we consider using the test data set to test the model to predict at a confidence level of 0.50, it is found that YOLOv5n and YOLOv8n can predict more accurately than YOLOv6n and YOLOv7, as shown in Figure 3.

(A) YOLOv5n

(B) YOLOv6n

(C) YOLOv7

(D) YOLOv8n

Fig. 3: Results obtained from experiments on the test dataset.

Based on the experimental results, we will focus our presentation on YOLOv5n and YOLOv8n as they are the models with the highest accuracy values. It can be seen that the confusion matrix results have similar classification values, with YOLOv5n being able to predict 90% of garbage images with 96% of bin images, while YOLOv8n can predict 96% of garbage images with accuracy and predict the accuracy of 91% of garbage images and 84% of bin images, Table 2 and Figure 4.

Table 2. Confusion matrix results of YOLOv5n and YOLOv8n

| True Label | | | |
|---|---|---|---|
| YOLOv5n | | | |
| Predicted | Garbage | Bin | Background |
| Garbage | 0.90 | 0.00 | 0.73 |
| Bin | 0.00 | 0.96 | 0.27 |
| Background | 0.10 | 0.04 | 0.00 |
| YOLOv8n | | | |
| Garbage | 0.91 | 0.00 | 0.72 |
| Bin | 0.00 | 0.84 | 0.28 |
| Background | 0.09 | 0.06 | 0.00 |

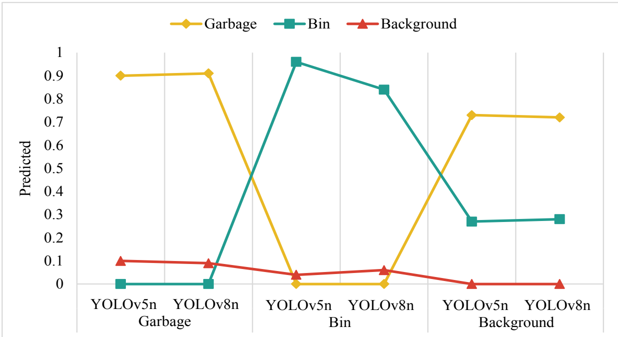



Fig. 4: mAP rate for comparison of YOLOv5n and YOLOv8n

We can apply the results from this experiment to develop a mobile or web application to notify of overflowing garbage bins as the resulting model can accurately classify or predict garbage and bins. The establishment of this system should be synchronized with CCTV cameras deployed in different communities throughout Bangkok, aiming to observe overflowing garbage bins and promptly alert relevant agencies or individuals in the vicinity. This integration facilitates quicker and more efficient responses, [21]. Although using a different version of YOLO, the goal is to focus on garbage detection in urban communities. This research applied a refined YOLOv3 network model utilized for garbage identification and classification. Based on the dataset gathered for this purpose, the network has been optimized. The findings indicate that the suggested strategy might make a significant contribution to smart cities' ability to manage garbage more effectively. Additionally, the study strives to enhance the precision of automatic garbage recycling, accelerate computational processes, and reduce the model size, all while ensuring practical applicability in real-world garbage recycling scenarios, [22]. This research presents a unique YOLO-based neural network model using Variational Autoencoder (VAE). A decoder, a convolutional predictor, and a convolutional feature extractor make up the model. After training, this model outperforms current models like YOLOv1 and Fast R-CNN, with a correct rate of 69.70% and 32.1 million parameters, processing at a pace of 60 frames per second (FPS). Moreover, the study affirms that using YOLO in garbage detection proves to be an effective means of controlling garbage pollution, [23]. This research primarily enhances the spatial pyramid pooling with average pooling, mish activation function, concatenated densely connected neural network, and hyperparameter optimization, this research provided an optimized YoLOv4-tiny model to detect floating junk. With a size of 16.4 MB and better results of 74.89% mean average precision, the suggested model is the best compromise among the other models. In terms of model size, detection time, and memory space, the suggested model performs well and may be integrated into low-cost devices.

## 5 Conclusion and Future Work

This research applied YOLO to detect 1,383 images of overflowing garbage bins, classified into 2 classes: garbage class and bin class. Currently, as YOLO has been developed in several versions, we compare each version, including YOLOv5n, YOLOv6n, YOLOv7, and YOLOv8n, using the smallest hyperparameter size as it has features to classify images. The comparison results showed that YOLOv5n had the highest mAP:0.5 value, and the model was able to classify with an accuracy of 94.50%, which when compared with YOLOv8n, the values were not much different. Using the test dataset to test the model to predict at a confidence level of 0.50, it is found that YOLOv5n and YOLOv8n can predict more accurately than YOLOv6n and YOLOv7. Considering the confusion matrix results, YOLOv5n can predict the accuracy of 90% of garbage images and 96% of bin images while YOLOv8n can predict 91% of garbage images and 84% of bin images.

Nonetheless, Bangkok or related organizations can utilize the model to expand, for instance, by developing a smartphone application that links to CCTV cameras and grants public officials and citizens access to the system. When the system notices that garbage bins are overflowing, it will promptly alert the local populace or relevant agencies. This is to allow people to take part and learn about the consequences of pollution from overflowing garbage bins. Additionally, it facilitates effective pollution reduction and timely waste management.

Therefore, we intend to create an alert system that will subsequently sound upon CCTV cameras detecting overfilled bins. Notifications will be sent to the appropriate authorities and the neighboring residents, along with an image and location of the overflowing garbage bins.

*Acknowledgement:*
This research was funded by the Faculty of Applied Science, King Mongkut's University of Technology North Bangkok. Contract no. 652115.

*References:*

[1] A. Singh, Managing the uncertainty problems of municipal solid waste disposal, *Journal of Environmental Management*, Vol.240, 2019, pp. 259-265.

[2] P. Nathapakti, N. Kulnides, and N. Supatanyaporn, Environmental problem, laws, consideration, and environmental case affecting the establishment of the environmental court in Phuket province, *Suan Sunandha Academic & Research Review*, Vol.13, No.2, 2020, pp. 1-12.

[3] W. Suppasedsak, Municipal solid waste problem and its solution with waste to energy power plant project, *The National Defence College of Thailand Journal*, Vol.61, No.3, 2019, pp. 107-119.

[4] R. M. Sebastian, D. Kumar, and B. J. Alappat, A technique to quantify incinerability of municipal solid waste, *Resources, Conservation & Recycling*, Vol.140, 2019, pp. 286-296.

[5] S. Srimeuang and P. Yolchayawong, Networking of participation in solving community waste problems in Nong Khaem district, Bangkok, *Journal of Graduate Schol, Pitchayatat*, Vol.15, No.1, 2020, pp. 203-211.

[6] P. Phodjanawichaikuland and J. Mekparyup, The study of the waste amount in Bangkok with cluster analysis, *Burapha Science Journal*, Vol.22, No.1, 2019, pp. 308-318.

[7] Environment Department Bangkok Metropolitan Administration, *Solid waste statistics, the fiscal year 2022*, [Online]. https://webportal.bangkok.go.th/public/user_files_editor/231/WasteData/2565/12.September2022.pdf (Accessed Date: April 5, 2023).

[8] H. Choi and J. Jeong, Realtime detection of table objects and text areas for OCR preprocessing, *WSEAS Transactions on Information Science and Applications*, Vol.20,

2023, pp. 197-205, https://doi.org/10.37394/23209.2023.20.23.
[9] K. Pathak, A. Arya, P. Hatti, V. Handragal, and K. Lee, A study of different disease detection and classification techniques using deep Learning for the cannabis plant, *International Journal of Computing and Digital Systems*, Vol.10, No.1, 2021, pp. 54-62.
[10] S. Kumar, M. Jailia, and S. Varshney, Improved YOLOv4 approach: A real time occluded vehicle detection, *International Journal of Computing and Digital Systems*, Vol.12, No.1, 2022, pp. 489-497.
[11] F. Mushtaq, K. Ramesh, S. Deshmukh, T. Ray, C. Parimi, P. Tandon, and P.K. Jha, Nuts&bolts: YOLO-v5 and image processing based component identification system, *Engineering Applications of Artificial Intelligence*, Vol.118, 2023.
[12] C. Gupta, N.S. Gill, P. Gulia, and J. Chatterjee, A novel finetuned YOLOv6 transfer learning model for real-time object detection, *Journal of Real-Time Image Processing*, Vol.20, 2023, pp. 1-19.
[13] J. Zhou, Y. Zhang, and J. Wang, A dragon fruit picking detection method based on YOLOv7 and PSP-Ellipse, *Intelligent Sensing and Machine Vision in Precision Agriculture*, Vol.23, No.8, 2023.
[14] R. Bawankule, V. Gaikwad, I. Kulkarni, S. Kulkarni, A. Jadhav, and N. Ranjan, Visual detection of waste using YOLOv8, *Proceedings of the International Conference on Sustainable Computing and Smart Systems (ICSCSS 2023)*, IEEE Xplore Part Number: CFP23DJ3-ART.
[15] F. Li, and L. Wang, Application of deep learning based on garbage image classification, *WSEAS Transactions on Computers*, Vol.21, 2022, pp. 277-282, https://doi.org/10.37394/23205.2022.21.34.
[16] Y. Liu, Y. Wang, Y. Li, Q. Li, and J. Wang, SETR-YOLOv5n: A lightweight low-light lane curvature detection method based on fractional-order fusion model. *IEEE Access*, Vol.10, 2022, pp. 93003-93016.
[17] C. Zheng, Stack-YOLO: A friendly-hardware real-time object detection algorithm, *IEEE Access*, Vol.11, 2023, pp. 62522-62534.
[18] L. Yang, Investigation of You Only Look Once Networks for Vision-based Small Object Detection, *International Journal of Advanced Computer Science and Applications (IJACSA)*, Vol.14, No.4, 2023, pp. 69-82.
[19] Z. Wang, Z. Hua, Y. Wen, S. Zhang, X. Xu, and H. Song, E-YOLO: Recognition of oestrus cow based on improved YOLOv8n model, *Expert Systems with Applications*, Vol.238, Part E, 15, 2024, pp. 1-17.
[20] A. Malta, M. Mendes, and T. Farinha, Augmented reality maintenance assistant using YOLOv5, *Applied Sciences*, Vol.11 No.11, 2021, pp. 475.
[21] B. D. Carolis, F. Ladogana, and N. Macchiarulo, YOLO TrashNet: garbage detection in video streams, *2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, Bari, Italy.
[22] A. Ye, B. Pang, Y. Jin, and J. Cui, A YOLO-based neural network with VAE for intelligent garbage detection and classification, *ACAI '20: Proceedings of the 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence*, December 2020 Article No.: 73, pp. 1-7.
[23] N. A. Zailan, A. S. M. Khairuddin, K. Hasikin, M. H. Junos, and U. Khairuddin, An automatic garbage detection using optimized YOLO model, *Signal, Image and Video Processing*, 2023.

**Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

The authors equally contributed to the present research, at all stages from the formulation of the problem to the final findings and solution.

**Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

This research was funded by the Faculty of Applied Science, King Mongkut's University of Technology North Bangkok. Contract no. 652115.

**Conflict of Interest**

The authors have no conflicts of interest to declare that are relevant to the content of this article.